



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Matemáticas

Escuela Profesional de Estadística

Análisis de árboles de clasificación para la creación de un modelo que permita describir el perfil de los morosos y no morosos en una entidad financiera de Lima Metropolitana en el 2014

TESINA

Para optar el Título Profesional de Licenciada en Estadística

AUTOR

Edelina Anyela TAFUR TEMBLADERA

ASESOR

Mg. Erwin KRAENAU ESPINAL

Lima, Perú

2016

Referencia bibliográfica

Tafur, E. (2016). *Análisis de árboles de clasificación para la creación de un modelo que permita describir el perfil de los morosos y no morosos en una entidad financiera de Lima Metropolitana en el 2014*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

Universidad del Perú, DECANA DE AMÉRICA

FACULTAD DE CIENCIAS MATEMÁTICAS

PROGRAMA DE ACTUALIZACIÓN PARA LA TITULACIÓN PROFESIONAL 2016 - II
MODALIDAD EXAMEN DE SUFICIENCIA PROFESIONAL
(RR. N° 03849-R-16)

Nº 147

ESCUELA PROFESIONAL DE ESTADÍSTICA

ACTA DE EXPOSICIÓN DE TESINA

✓
11-R
54-P

En la Ciudad Universitaria, Facultad de Ciencias Matemáticas, siendo las ~~13:00~~ horas, del día 11 de diciembre del 2016, se reunieron las docentes designadas como miembros del Jurado Evaluador:

- | | |
|--------------------------------------|------------|
| - Dr. Luis Alberto Navarro Huamani | Presidente |
| - Mg. Manuel Rolando Canales del Mar | Miembro |

Para la exposición de la Tesina titulada: «ANÁLISIS DE ARBOLES DE CLASIFICACIÓN PARA LA CREACIÓN DE UN MODELO QUE PERMITA DESCRIBIR EL PERFIL DE LOS MOROSOS Y NO MOROSOS EN UNA ENTIDAD FINANCIERA DE LIMA METROPOLITANA EN EL 2014», presentada por la Bachiller **Edelina Anyela Tafur Tembladera**.

Luego de la exposición de la tesina, los Miembros del Jurado hicieron las preguntas correspondientes, a las cuales la Bachiller Edelina Anyela Tafur Tembladera respondió con acierto y solvencia, demostrando pleno conocimiento del tema.

Hecha la evaluación correspondiente, según tabla adjunta, la Bachiller Edelina Anyela Tafur Tembladera mereció la aprobación obteniendo como calificativo promedio y la nota de Dieciocho (18) (letras y números).

A continuación los Miembros del Jurado Evaluador, dan manifiesto que la Bachiller **Edelina Anyela Tafur Tembladera**, aprobó la exposición de la Tesina.

Siendo las ~~13:00~~ horas, se levantó la sesión, firmando para constancia la presente acta en dos (2) copias originales.

Mg. Manuel Rolando Canales del Mar
MIEMBRO

Dr. Luis Alberto Navarro Huamani
PRESIDENTE

RESUMEN

ANÁLISIS DE ÁRBOLES DE CLASIFICACIÓN PARA LA CREACIÓN DE UN MODELO QUE PERMITA DESCRIBIR EL PERFIL DE LOS MOROSOS Y NO MOROSOS EN UNA ENTIDAD FINANCIERA DE LIMA METROPOLITANA EN EL 2014

BR. Edeline Anyela Tafur Tembladera

Diciembre 2016

ASESOR : Mg. Erwin Kraenau Espinal

TÍTULO OBTENIDO : Licenciada en Estadística

Las entidades financieras están desarrollando constantemente modelos que les ayuden a predecir el comportamiento de sus clientes, con el fin de mejorar sus ganancias o de reducir las pérdidas, como el riesgo de “no pago” o de “morosidad”.

El presente trabajo busca crear un modelo predictivo de Árboles de Decisión por Clasificación con minería de datos para predecir la morosidad financiera en Lima Metropolitana, describiendo y analizando las variables independientes con el algoritmo CHAID (Detector automático de Chi-cuadrado de interacción) siguiendo la metodología CRISP-DM en una base de datos de una entidad financiera. La principal ventaja de esta metodología es que no está sujeta a supuestos estadísticos, este método es muy usado cuando se busca predecir o clasificar comportamientos de riesgo, como la morosidad, y se cuenta con más variables cualitativas que cuantitativas. La metodología persigue la separación óptima en la muestra, de tal modo que los grupos de la variable respuesta ofrecen distintos perfiles de riesgo.

***PALABRAS CLAVES: Árboles de Clasificación CHAID, Minería Datos,
perfiles de morosidad***

ABSTRACT
**ANALYSIS OF CLASSIFICATION TREES FOR THE
CREATION OF A MODEL THAT ALLOWS TO DESCRIBE
THE PROFILE OF THE NON-MOROUS AND NON-LOAN
IN A FINANCIAL ENTITY OF LIMA METROPOLITANA IN
2014**

BR. Edelina Anyela Tafur Tembladera
Diciembre 2016

ADVISOR : Mg. Erwin Kraenau Espinal
DEGREE OBTAINED : Bachelor`s degree in Statistics

Financial institutions are constantly developing models that help them predict their clients' behavior in order to improve their profits or reduce losses, such as the risk of "non-payment" or "delinquency." The present work seeks to create a predictive model of Decision Trees by Classification with data mining to predict financial delinquency in Metropolitan Lima, describing and analyzing the independent variables with the algorithm CHAID (Automatic Chi-square interaction detector) following the methodology CRISP-DM in a database of a financial institution. The main advantage of this methodology is that it is not subject to statistical assumptions; this method is widely used when it is sought to predict or classify risk behaviors, such as delinquency, and there are more qualitative than quantitative variables. The methodology pursues the optimal separation in the sample, so that the groups of the response variable offer different risk profiles.

KEYWORDS: CHAID Classification Trees, Mining Data, delinquency profiles.